

Glycoproteomics-Compatible MS/MS-Based Quantification of Glycopeptide Isomers

Joshua C.L. Maliepaard, J. Mirjam A. Damen, Geert-Jan P.H. Boons, and Karli R. Reiding*

Cite This: *Anal. Chem.* 2023, 95, 9605–9614

Read Online

ACCESS |



Metrics & More

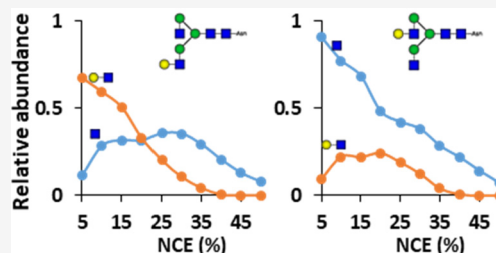


Article Recommendations



Supporting Information

ABSTRACT: Glycosylation is an essential protein modification occurring on the majority of extracellular human proteins, with mass spectrometry (MS) being an indispensable tool for its analysis, that not only determines glycan compositions, but also the position of the glycan at specific sites via glycoproteomics. However, glycans are complex branching structures with monosaccharides interconnected in a variety of biologically relevant linkages, isomeric properties that are invisible when the readout is mass alone. Here, we developed an LC-MS/MS-based workflow for determining glycopeptide isomer ratios. Making use of isomerically defined glyco(peptide) standards, we observed marked differences in fragmentation behavior between isomer pairs when subjected to collision energy gradients, specifically in terms of the galactosylation/sialylation branching and linkage. These behaviors were developed into component variables that allowed for relative quantification of isomerism within mixtures. Importantly, at least for small peptides, the isomer quantification appeared to be largely independent from the peptide portion of the conjugate, allowing a broad application of the method.



INTRODUCTION

Glycosylation is a ubiquitous form of post-translational modification (PTM) that occurs on the majority of extracellular proteins.^{1,2} Protein N-glycosylation, glycosylation of the side-chain of an Asn residue, primarily takes place within an Asn-Xxx-Ser/Thr motif (Xxx ≠ Pro) and often on multiple sites within the same protein.² Glycans themselves consist of monosaccharides interconnected via glycosidic bonds; typical monosaccharides include galactose (Gal), mannose (Man), and glucose (Glc), collectively referred to as hexoses (Hex) due to their shared six-ring and mass, N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc), referred to as N-acetylhexosamines (HexNAc), all in addition to fucoses (Fuc) and sialic acids of which N-acetylneuraminic acid (NeuAc) is the primary human variant.³ The carbon with which one monosaccharide connects to another determines the terminology of the glycan linkage; for each pair of monosaccharides, a set of numbers describes which carbons are joined via O-glycosidic linkage, e.g., 1–4, whereas the addition of the prefix α or β informs on the anomers of the bond.¹ A typical example of this is lactose, in which a Gal is β 1–4-linked to a Glc, i.e., Gal β -(1–4)-Glc.¹ The wealth of possible linkages between monosaccharides and the resulting branching glycan structures represent a large reservoir of biological variation between one glycoproteoform and the next.

Glycans are involved in a variety of cellular processes, and even a slight change in the structure of a glycan can have a major impact on the function of a glycoprotein or the way it interacts with its environment. One example is the way in which sialic acid linkage effects the ability of influenza viruses

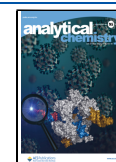
like the H5N1 avian influenza to cross the species barrier.⁴ Avian influenza viruses mostly infect via binding α 2,3-linked sialylation, which lines the avian respiratory track, while the human tract is mostly lined with α 2,6-linked sialylation.⁵ However, swine carry both α 2,3- and α 2,6-linked sialic acids, making them ideal candidates for cross-species contamination and subsequent mutation.⁵

Mass spectrometry (MS) is a powerful tool to perform glycosylation analysis—the mass of a glycan, glycopeptide, or glycoprotein is highly informative for acquiring the composition of glycosylation.⁶ However, glycan isomerism is extremely challenging to study from mass alone. For released glycan analysis, so without information on the glycosylation site, certain isomeric aspects can already be determined by the separation method prior to the MS analysis. For instance, porous graphitized carbon liquid chromatography (PGC-LC) can separate α 2,3-/ α 2,6-sialylation and β 1,3-/ β 1,4-galactosylation,^{7,8} hydrophilic interaction liquid chromatography (HILIC) allows distinction between sialic acid linkages and galactose branching,^{9,10} while capillary electrophoresis has been instrumental in separating the sialic acid linkages as well.¹¹ Next to this, chemical derivatization of sialic acids is an

Received: March 27, 2023

Accepted: June 5, 2023

Published: June 15, 2023



attractive strategy to add a mass tag to different linkages,¹² collision-induced dissociation can provide isomerism-dependent fragments and fragment ratios,^{13,14} and ion mobility (IM)-MS is rapidly maturing as a gas-phase separation method for analyzing isomerism as well.^{15–17} However, despite extensive development in determining the isomerism of released glycans, work on glycopeptides remains very sparse. This class of analytes can provide site-specific information, essential for multiplying glycosylated proteins or protein mixtures, but the peptide portion presents a major analytical challenge. Sialic acid linkages appear accessible from an MS² or MS³ level, although its assessment was primarily performed for O-glycopeptides.¹⁴ Recent studies have shown the ability of IM-MS to separate sialic acid linkage isomers;^{18–20} however, these methods have not yet shown the ability to distinguish between other isomeric properties such as branching isomerism. IM-MS also necessitates the use of IM-capable instruments, whereas a MS²-based approach would be applicable to a wider range of instruments. As it is, there currently appears to be no proteomics-compatible method to determine glycan structural isomerism for a broad range of isomeric properties, let alone the ratios thereof for a given glycosylation site.

As such, we present here a proteomics-compatible method for the relative quantification of isomer ratios at the glycopeptide level, suitable for the determination and relative quantification of sialic acid linkages and sialic acid/galactose branching positions on N-glycopeptides. We have developed this method by using isomerically defined glycosylated asparagine standards, generated by chemoenzymatic synthesis,²¹ and show broad applicability by analyzing isomer mixtures and differing peptide lengths. Importantly, the determination of glycopeptide isomerism only requires MS/MS, namely, a cycle through collision energies, and can therefore easily be applied without using special instrumentation or modes of fragmentation.

METHODS

Chemicals and Reagents. Milli-Q water was generated from a Merck Milli-Q IQ 7003 system (Darmstadt, DE). Pronase and trastuzumab were obtained from Roche (Woerden, NL). α -(1-2,3,4,6)-L-Fucosidase was obtained from Megazyme (Ayr, U.K.). Chloroacetamide (CAA), Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), Tris, trypsin, Lys-C, sodium deoxycholate (SDC), sodium acetate, and formic acid (FA) were obtained from Merck (Darmstadt, DE). Acetonitrile and 0.1% FA were obtained from Biosolve (Valkenswaard, NL). Egg yolk powder was obtained by freeze-drying egg yolks from eggs obtained from SPAR (Waalwijk, NL). Cotton string was obtained from Kruidvat (Renswoude, NL). Trifluoroacetic acid (TFA) was obtained from Honeywell International Inc. (Charlotte, NC).

Glycosylated Asparagine Standards. Isomerically defined glycosylated asparagine standards were synthesized and characterized as previously described.^{22–24} The standards varied in either galactose positioning (3-branch and 6-branch), sialic acid positioning (3-branch and 6-branch, one galactose and two), or sialic acid linkage (α 2,3- and α 2,6-linked). Mixes of each isomer pair were made ranging from 100% isomer A and 0% isomer B to 0% isomer A and 100% isomer B in steps of 10%.

Glycopeptides Digestion and Purification. Sialylglycopeptide (SGP) and trastuzumab were used as model proteins for linkage determination. SGP was purified from egg yolk

powder using HILIC-SPE. In short, 500 μ g of egg yolk powder was dissolved in 250 μ L of 80% ACN/0.1% TFA. A total of 1 cm of cotton string was used as column material, as previously described.²⁵ The column was conditioned 3 times with 200 μ L of 0.5% TFA in Milli-Q, then washed with 3 times with 200 μ L of 80% ACN/0.5% TFA. After washing, the sample was loaded into the column 10 times. After sample loading, the column was washed again with 3 times with 200 μ L of 80% ACN/0.5% TFA and subsequently eluted 2 times with 200 μ L of 50% ACN/0.5% TFA. The eluent was dried in a vacuum concentrator and subsequently brought to 100 μ L with 0.1 M Tris and 10 mM CaCl₂ for Pronase digestion.

The Pronase digestion was performed at 37 °C and stopped at 7.5, 15, 30, 60, 120, and 240 min, and after overnight digestion, to obtain a range of glycopeptides with varying peptide length. Trastuzumab was first reduced and alkylated in a pH 8.5 100 mM Tris, 10 mM TCEP, and 40 mM chloroacetamide 1% SDC buffer for 30 min at room temperature. Subsequently, reduced and alkylated trastuzumab was digested overnight using trypsin and Lys-C in 50 mM ammonium bicarbonate at room temperature. Half of the tryptic trastuzumab digest was dissolved in 100 mM NaAc pH 4.5 and defucosylated using α -(1-2,3,4,6)-L-fucosidase overnight at 37 °C, half of which was enriched by HILIC-SPE. All fractions were digested with Pronase for 15, 30, 60, 120 min and overnight.

LC-MS Method. Measurements were performed on Thermo scientific Exploris 480 and Fusion systems connected to Thermo Scientific Ultimate 3000 UHPLC systems. A 50 cm, 75 μ m ID, 2.4 μ m repositil column was used with a 60 min gradient of 2% B at 0–1 min, 13% B at 13 min, 44% B at 42 min, 99% B at 44–49 min, and 2% B 50–60 min. Source voltage was set to 2100 V. Maximum time between MS scans was set to 3 s, allowing for 10 data points on a standard 30 s-wide LC-peak. After an initial MS scan at scan range m/z 375–2000, resolution 120000, automatic gain control (AGC) set to standard, ions with a 2+ charge and an intensity of over 5×10^3 were compared to a precursor list (Figure 3), which was used to trigger a MS/MS scan using higher-energy collisional dissociation (HCD) at 29% normalized collision energy (NCE) at a scan range of m/z 120–4000, resolution 60000, AGC target set to standard, and a maximum injection time of 20 ms. If at least 3 oxonium ions from the mass trigger list (Figure S1) were detected, 10 subsequent MS/MS scans were performed of the same precursor using NCEs 5–50%. Since the mass trigger list covered more than 3 HexNAc-derived oxonium ions, including a HexNAc (m/z 204.0867) and a HexNAc with 1 or 2 water losses (m/z 186.0761 and m/z 168.0655, respectively), every glycopeptide with a complex N-glycan was expected to trigger a NCE cycle. To minimize any potential influence of the scan order on results, a semirandom scan order was used (NCEs: 30%, 10%, 50%, 20%, 40%, 15%, 45%, 25%, 35%, and 5%), and the order was repeated in triplicate. After the MS/MS scans, the precursor was excluded from triggering further MS/MS scans for 15 s. To determine the optimal charge state, the method described above differed by not only covering 2+, but rather all charge states between 2+ and 8+.

Data Analysis. Thermo raw files were converted to MGF format using Proteowizard MSconvert (version 3.0.21328–404bcbf1) using MGF as output format, 64-bit binary encoding precision, and the following options selected: write index, zlib compression, and TPP compatibility. MGF files were searched

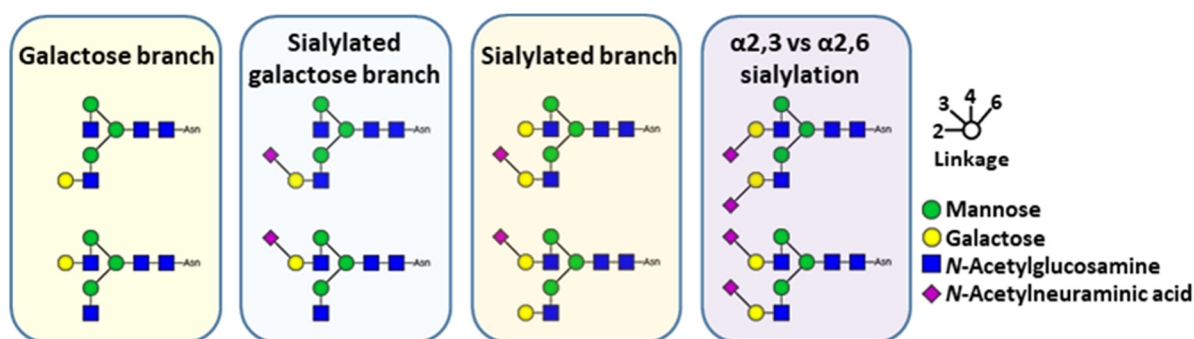


Figure 1. Overview of isomerically defined glycosylated asparagine standards. A total of eight isomerically defined glycopeptides were combined into four pairs of linkage characteristics. Three pairs covered the branching position, having antennary extension on either the α 1,3- or α 1,6-linked Man, allowing the investigation of Gal-, NeuAcGal-, and NeuAc-branching. The other glycopeptide pair differed in the sialic acid linkage, allowing the investigation of α 2,3- and α 2,6-sialylation.

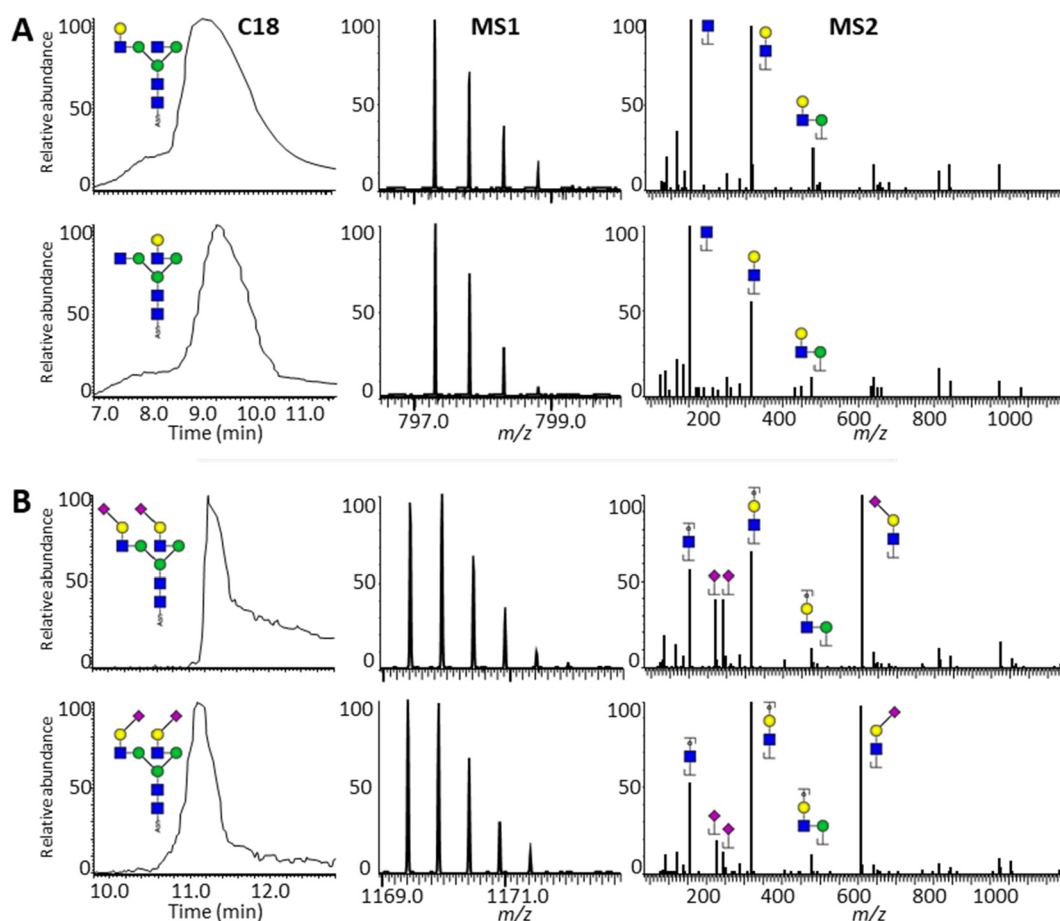


Figure 2. Comparison of N4H4 and N4H5S2 isomers in C18 chromatography, MS, and MS/MS. (A) Separation of galactose branching isomers by C18 chromatography (left). Signal of the $[M + 2H]^{2+}$ ions corresponding to the galactose branching isomer species. (middle). MS/MS spectra were generated by the galactose branching isomers when subjected to HCD at NCE 20% (right). (B) Separation of sialic acid linkage isomers by C18 chromatography (left). Signal of the $[M + 2H]^{2+}$ ions corresponding to the sialic acid linkage isomer species. (middle). MS/MS spectra were generated by the sialic acid linkage isomers when subjected to HCD at NCE 20% (right). As can be seen, while LC and MS do not show differences between the isomers, MS/MS shows discernible peak ratio differences.

for spectra containing glycan oxonium ions using an internally developed tool named PeakSuite (v1.10.1) using an ion delta of 20 ppm, noise filter of 0%, and using a list of oxonium ion m/z values as mass targets (Table S1). Python 3.2.2 was then used to (1) assign collision energies to the spectra, (2) for data curation based on retention time, precursor m/z , and oxonium ion intensity, and (3) for calculating the relative intensity of

the oxonium ions (the Python script can be found as Supporting Information, 1). The relative intensities of the oxonium ions were calculated by dividing the intensity of an oxonium ion by the summed intensities of all oxonium ions in the same scan.

For construction of linkage variables, the following curation criteria were used: a 10–20 min retention time window for

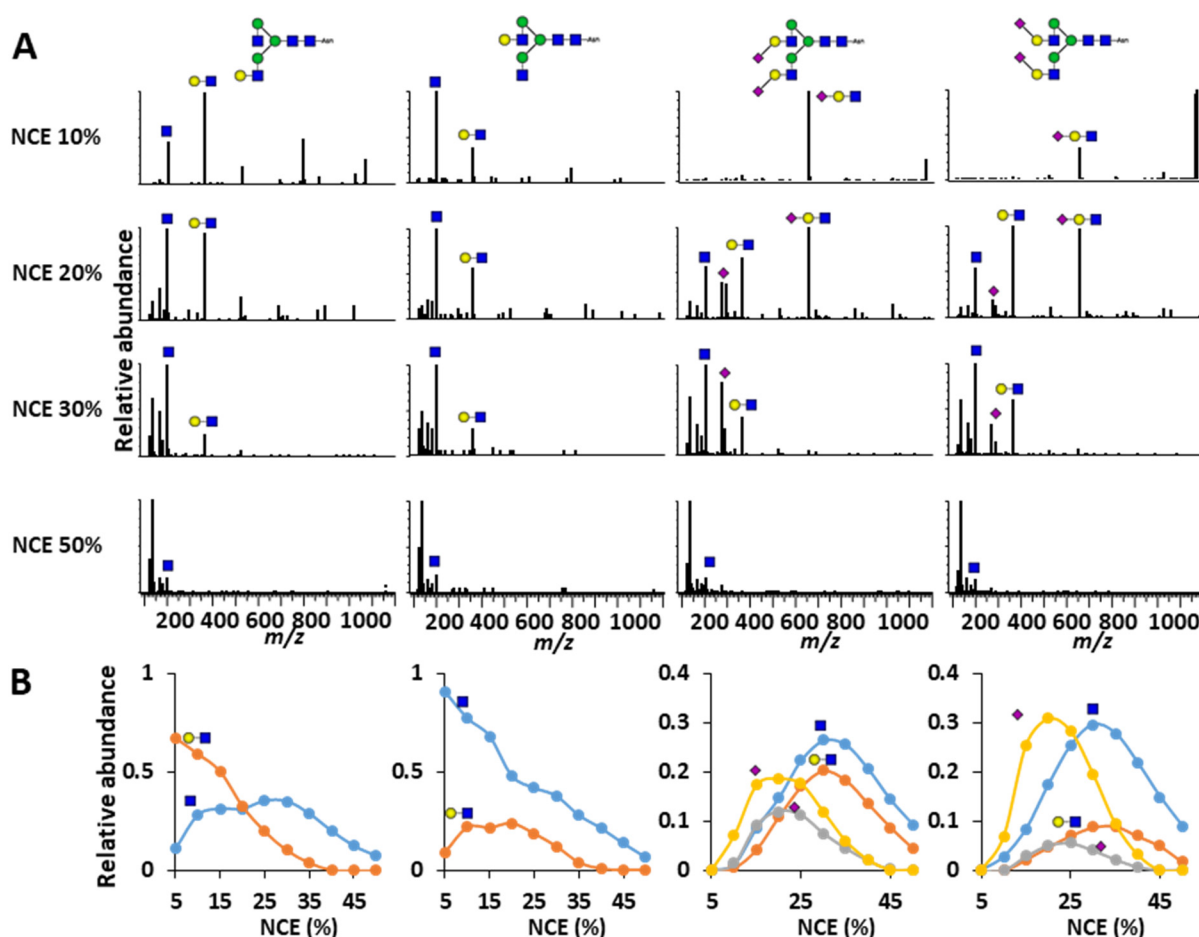


Figure 3. MS/MS differences between glycopeptide isomers. (A) MS/MS at NCEs of 10, 20, 30 and 50% of the N4H4 glycan with galactosylation positioned on the 3-branch (full left) or the 6-branch (center left) and the N4H5S2 glycan with full α 2,3-linkage (center right) or full α 2,6-linkage (full right). (B) Relative intensity differences between the oxonium ions HexNAc and HexHexNAc for 5–50% NCE. For all galactosylation and sialylation isomers, NCE regions could be determined in which the isomers were distinguishable. For galactosylation isomers, the NCE range was 10–25%, while for sialylation isomers, the NCE range was 20–45%.

samples run with a 60 min gradient, a 5–20 min retention time window for samples run with a 30 min gradient, a 5×10^3 intensity threshold for NeuAc B–H₂O (m/z 274.0921) and a 1×10^4 intensity threshold for HexNAc, HexHexNAc, and NeuAcHexHexNAc B ions (m/z 204.0867, 366.1395, 657.2349, respectively) in the samples in which they occurred. This data curation led to the elimination of MS/MS spectra that would otherwise be wrongfully included in the variable, such as apparent alternative glycopeptide compositions (Figure S2).

The curated data was analyzed using RStudio (2022.02.2 + 485) (Supporting Information, 2). For the glycosylated asparagine standards, level plots were made for each of the aforementioned oxonium ions with the NCE on the x -axis (seq(5,50,1)), the percentage of isomer A on the y -axis (seq(0,100,2)), and the relative intensity of the oxonium ions for the z -axis. On the basis of these outcomes, a variable was constructed that would report on the isomer ratios in each of the pair mixtures. For a full overview of linkage variable calculations, see Table S2. Relevant variables were subsequently used to report the isomeric properties of SGP and trastuzumab.

Statistical Information. Differences in linkage variables were tested by a two-tailed student's t test (Table S3). Results

were deemed significant at p -value ≤ 0.05 , with a false discovery rate of 5%.

Data Representation. Glycan cartoons were generated using Glycoworkbench 2.1 following the recommendations of the Consortium for Functional Glycomics.^{26,27}

RESULTS

Standard Identities. For MS determination of isomerism, we made use of eight isomerically defined glycosylated asparagine standards, as synthesized and characterized previously.^{22–24} Combined, these standards covered four distinct isomeric properties (Figure 1); three out of the four pairs covered branching, i.e., having an antennary extension on either the α 1,3- or α 1,6-linked Man allowing the investigation of Gal-, NeuAcGal-, and NeuAc-branching, whereas the fourth isomer pair differed in the linkage of the NeuAc, i.e., full occupation of either α 2,3- or α 2,6-linked NeuAc. Accordingly, the respective compositions of the standards were N4H4, N4H4S1, N4H5S1, and N4H5S2 (N = N-acetylhexosamine, H = hexose, and S = N-acetylneuraminic acid).

Fragmentation Behavior across the NCE Range. The isomerically defined glycosylated asparagine standards were analyzed by LC-MS/MS-based glycoproteomics with varying normalized collision energies (NCEs). Expectedly, the chromatography of the isomer pairs exhibited very similar

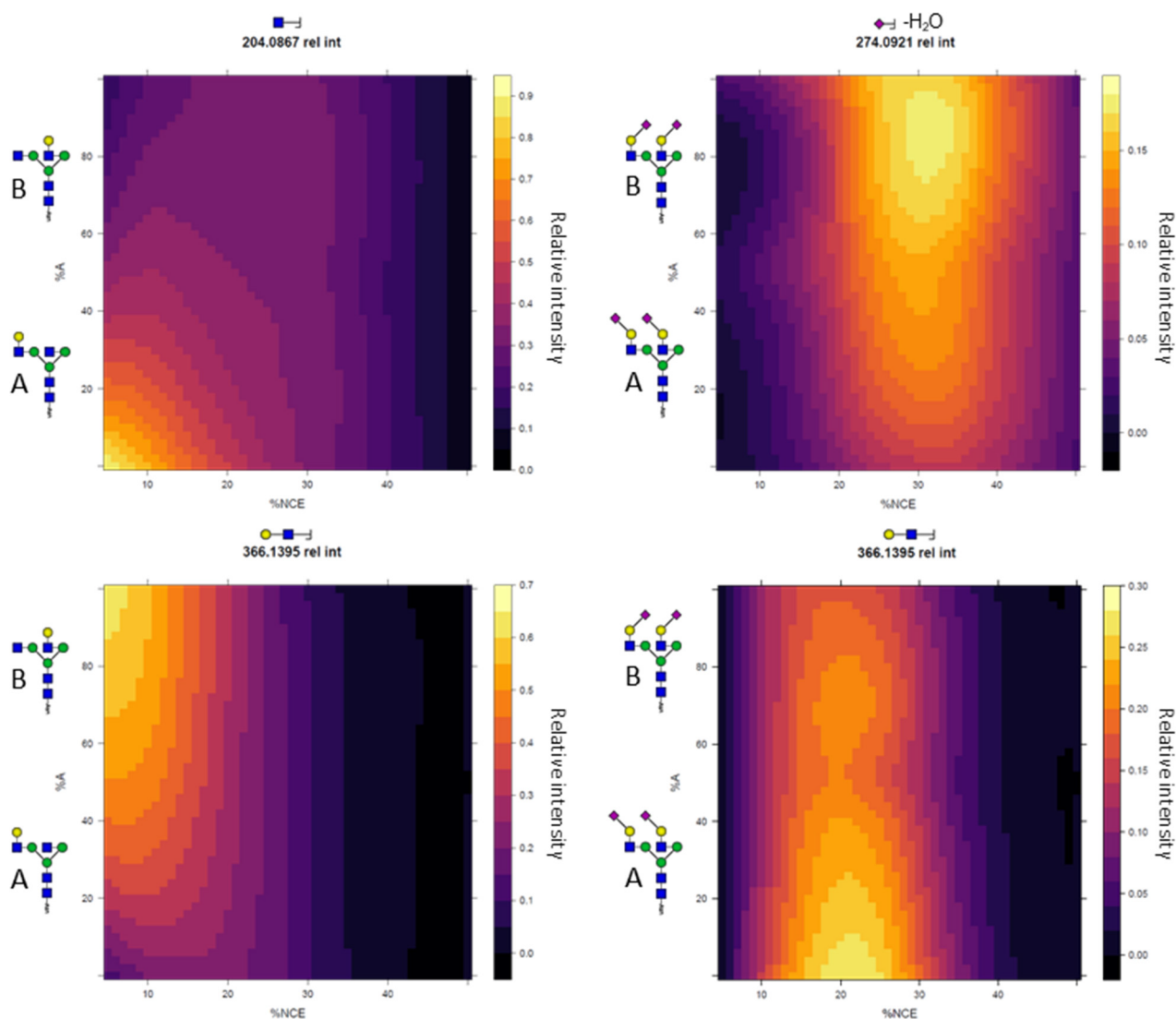


Figure 4. Oxonium ion intensity differences across isomer mixtures. Each panel represents a single oxonium ion. Displayed on the *x*-axis is the collision energy gradient in % NCE, on the *y*-axis is the ratio of the isomer pair in %A (left isomer), and on the *z*-axis (the color) is the intensity of the oxonium ion, normalized against all others within the same MS-run. As can be seen, not only do the presented oxonium ions show differences between isomers across the energy gradient, but they gradually changed across the isomer mixture as well.

peak shapes and retention times alongside an almost identical isotopic distribution in MS (Figure 2A,B). At the same time, clear differences in oxonium ion intensities could be observed between the isomer pairs when subjected to MS/MS at given NCEs, for example, between α 2,3- and α 2,6-sialylated glycosylated asparagine standards at 20% NCE (Figure 2B). Here, the relative abundances for *N*-acetylneuraminic acid (NeuAc) B- and B-H₂O-ions (*m/z* 292.1027 and 274.0921, respectively) both proved approximately 40% of the base peak intensity (BPI) for the α 2,3-linked variant, while these were only 15 and 20% for the α 2,6-linked variant. Similarly, B-ions for HexNAc (*m/z* 204.0866) and HexHexNAc (*m/z* 366.1395) showed BPIs of 58 and 68% for the α 2,3-sialylated glycopeptide standard and 54 and 100% for its α 2,6-sialylated equivalent. The differences in MS/MS were not constrained to the sialic acid linkages but could be observed for branching positions as well (Figure S3). For example, at 10% NCE, 3-branch monogalactosylated glycopeptides showed for HexNAc and HexHexNAc B-ions abundances of 44 and 100% BPI,

respectively, whereas the variant with the galactose on the 6-branch showed abundances of 100 and 28% instead.

These differences led us to investigate the behavior of a wider group of ions across the collision energy space (Figure S1). In order to maximize the amount of MS/MS scans of glycopeptide, a mass trigger was introduced: when at least 3 of oxonium ions from a list of 15 were detected in a preliminary MS/MS scan at 29% NCE, 10 subsequent MS/MS scans of the same precursor were performed at NCEs ranging from 5 to 50%. The NCE range was set at 5–50% NCE since NCE below 5% showed almost no fragmentation of the parent ion, while at NCE 50% the spectrum was dominated by *m/z* 138.0550 (B-ion for HexNAc –CH₆O₃) with very few other oxonium ions present. A step size of 5% NCE was chosen in order to keep the total scan cycle time in line with other proteomics methods (~250 ms) while still giving good coverage of the entire NCE range. Obtained intensities of the oxonium ions were normalized to the total intensity, and compared between isomer pairs (Figure 3). Only oxonium ions (B-ions) were taken into account, as opposed to Y-ions, since

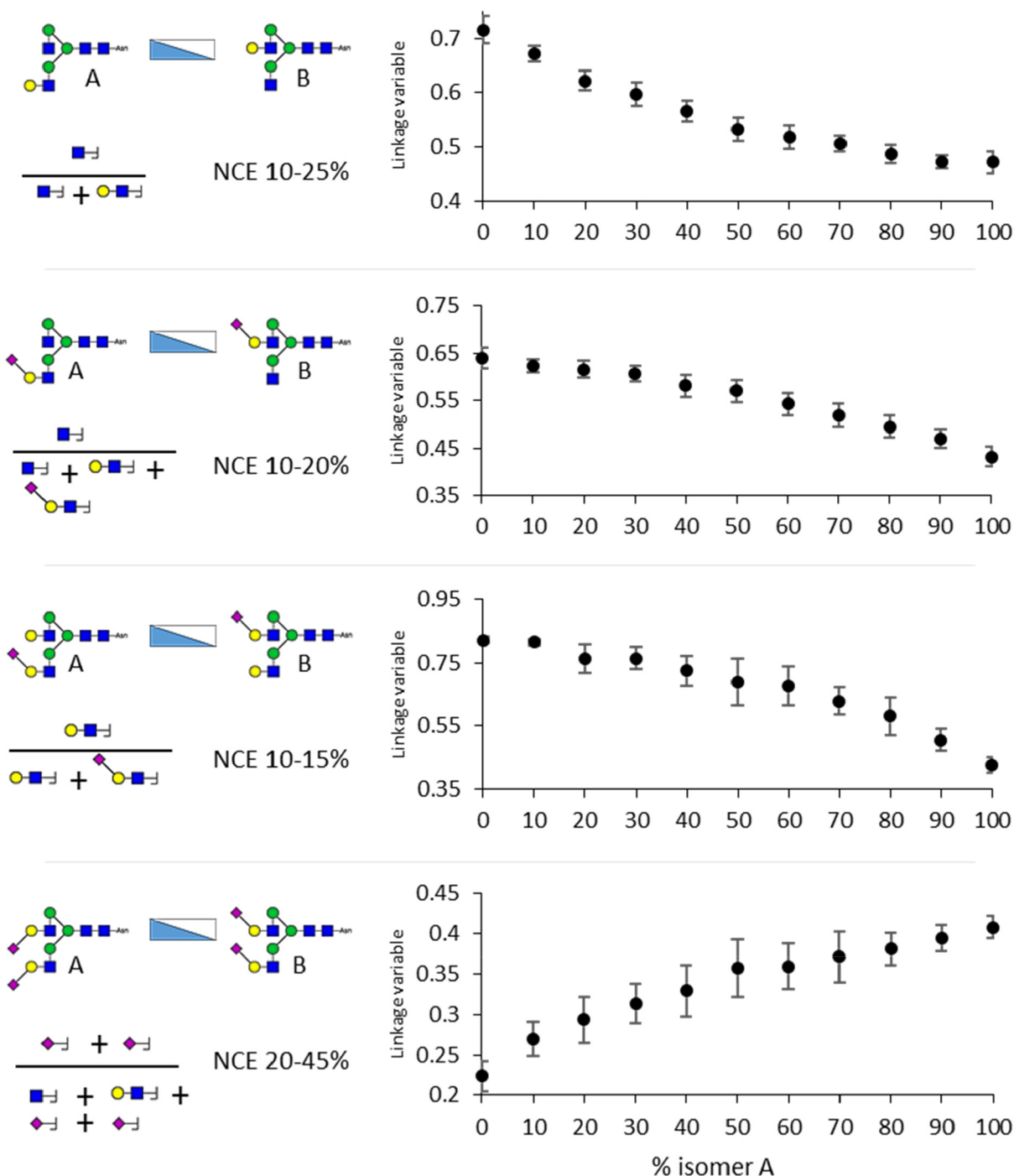


Figure 5. Average linkage variables for glycopeptide isomer mixes with standard deviations. The linkage variables for the glycopeptide isomers were based on the relative intensity of glycan oxonium ions across a range of collision energies, as indicated on the left. Already at differences of 10%, most mixtures differed with statistical significance. All dots represent the average \pm standard deviation. Linkage variables used for this figure were calculated on a per NCE-cycle basis.

these can principally be formed irrespective of the connecting peptide, thereby making translation to alternative peptide backbones more likely.

To determine which charge states would be most beneficial for distinguishing the glycopeptide isomers, the standard panel was primarily measured with a stepped-HCD method that allowed for fragmentation of analyte charge states ranging from 2+ to 8+. Interestingly, the main charge states observed for the standards, namely, 2+ and 3+, showed noticeable fragmentation differences, with 2+ having much more pronounced

relative intensity differences between isomer pairs than 3+ (Figure S4). One example of this was the relative intensity of HexNAc B (m/z 204.0867), in which 10% NCE in 2+ was 77% BPI for 6-branched galactosylation and 28% BPI for its isomeric counterpart (a 49% difference), whereas in 3+ these values were 48% BPI and 43% BPI (a 5% difference). For NeuAc B-H₂O (m/z 274.0921) at 30% NCE in 2+, the relative intensity was 20% BPI for α 2,3-linked sialylation and 9% BPI for its isomer (a 11% difference), whereas in 3+ these numbers were 22% BPI and 15% BPI (a 7% difference; Figure

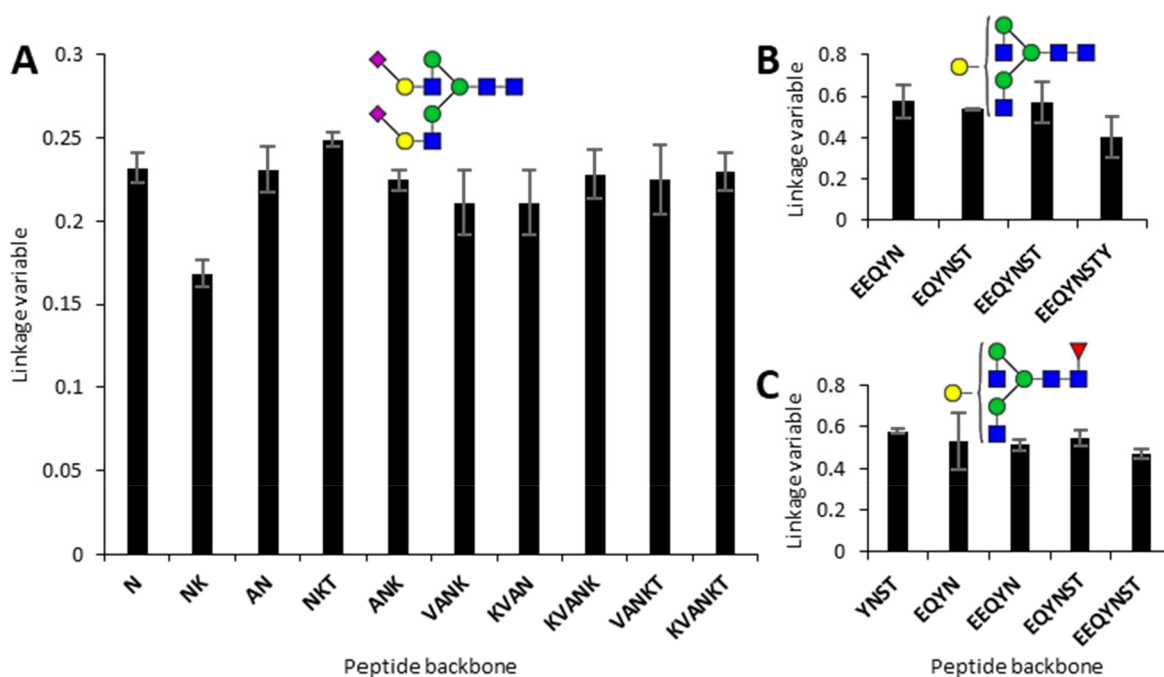


Figure 6. Effect of peptide backbone on sialic acid linkage and galactose branching assignment. (A) Partially digested SGP carrying N4H5S2. The linkage variable ranged from 0.2 to 0.25 across the glycopeptides, with the average value of NKT being highest at 0.25 and lowest at 0.17. These linkage variables corresponded with almost exclusively (90–100%) α 2,6-linked sialic acids, in accordance with the literature for SGP.²⁸ (B) Partially digested trastuzumab carrying N4H4. The linkage variable ranged from 0.4 to 0.57 across the glycopeptides with an average value of 0.54. This linkage variable corresponds to an approximate of 40–50% galactosylation of the 3-branch. (C) Partially digested trastuzumab with N4H4F1. The linkage variable ranged from 0.47 to 0.58 across the glycopeptides with an average value of 0.51. This linkage variable corresponds to approximately 60–70% galactosylation of the 3 branches. All bars represent the average \pm standard deviation. Linkage variables used for this figure were calculated on a per NCE-cycle basis.

S4). Because of this, the decision was made to proceed with measurement and data analysis of analytes with charge state 2+ specifically.

Construction of Variable for Isomer Determination.

Since clear differences were observed in the fragmentation behavior of the isomeric standards across collision energy, the next step was to attempt relative quantification within mixtures thereof. For this purpose, isomerically defined glycosylated asparagine standards were mixed ranging from 100% of one isomer (A) to 100% of the other (B) in steps of 10%. After LC-MS measurement, for the relative intensity for (1) each oxonium ion at (2) each measured collision energy across (3) each glycopeptide mixture, was interpreted for the purpose of isomer discrimination (Figures 4 and S5–S8).

Taking the monogalactosylated glycopeptide mixtures as an example, at an NCE of 10%, the relative abundance of the HexNAc B-ion (m/z 204.0867) increased from 20% base peak intensity (BPI) in the sample with the galactose on the 3-branch, to 80% BPI in the samples with the galactose on the 6-branch. At the same time, the relative abundance of the HexHexNAc B-ion (m/z 366.1395) increased from 25% to 60% BPI. In general, the differences in relative oxonium ion abundance were clearest at collision energies ranging from 10 to 25% NCE.

Since the differences in HCD fragmentation between isomers occurred over a range of collision energies and for multiple oxonium ions, a variable was created to condense these aspects into a single number.

For the monogalactosylated glycopeptides, the oxonium ions that showed the greatest difference were in the B-ions of HexNAc (m/z 204.0867) and HexHexNAc (m/z 366.1395),

most appreciably so between NCE values of 10 and 25% (Figure S5). Below NCE 10% the ion intensity appeared to be too variable, whereas above 25% the isomers no longer showed a difference. For variable construction, then, the relative abundance of HexNAc (m/z 204.0867) was divided by the sum of the relative abundance of the HexHexNAc and HexNAc B-ions (m/z 366.1395 and 204.0867, respectively), averaged across NCE 10–25% (i.e., divided by the number of included collision energies to yield a number between zero and one). The same process was followed for the sialylated and sialylated galactose glycopeptide isomers: for N4H4S1 both HexHexNAc and HexHexNAcNeuAc B-ions (respectively m/z 366.1395 and 657.2349) showed the largest difference across NCEs 10–15% (Figure S6), making the formula for linkage determination $366.1395/(366.1395 + 657.2349)/2$, whereas for N4H5S1 the ions HexNAc (m/z 204.0867), HexHexNAc (m/z 366.1395) and HexHexNAcNeuAc (m/z 657.2349) (Figure S7) led up to a linkage variable formula of $204.0867/(204.0867 + 366.1395 + 657.2349)/3$ between NCEs 10–20%. Finally, the difference in α 2,3- and α 2,6-sialylation for N4H5S2 proved most determinable by $(274.0921 + 292.1027)/(274.0921 + 292.1027 + 204.0921 + 366.1395)$ averaged across NCE 20–45% (Figure S8).

Across the isomer mixtures, the linkage variables showed clear statistically significant differences between the mixture percentage steps (Figure 5). For example, a two-tailed students t test showed a p -value < 0.05 for 0–10%, 10–20%, 20–30%, 40–50%, and 90–100% α 2,3-linkage with 1.64×10^{-7} when comparing 0 and 10% α 2,3-linkage, and 0.048 when comparing 20% and 30% α 2,3-linkage as examples. Also comparisons for galactose branching, with the exception of 60–70% and 90–

100%, consistently showed *p*-values below 0.01, examples being $p = 1.99 \times 10^{-6}$ for 30–40% and $p = 4.64 \times 10^{-4}$ for 70–80%. This allowed us to quantify the glycan structures of most of the tested glycopeptide mixtures with an approximate accuracy of 10%.

Glycopeptide Analysis. Subsequently, the effect of the peptide backbone on the linkage variable was tested to determine the applicability of the linkage variables for glycopeptides with different peptide backbones but the same glycan.

Sialylglycopeptide (SGP), a well-studied glycopeptide with a known glycan structure,²⁸ was used to assess whether the length of the connected peptide impacted fragmentation characteristics and whether this could be overcome by the linkage variable. To obtain a range of different peptides with the same glycan, we partially digested SGP with Pronase. This yielded a range of glycopeptides of different sizes, i.e., N (identical to the standard panel), AN, NKT, KVAN/VANK (same mass), VANKT, and KVANKT (the full peptide; Figure S9). Using the same LC-MS/MS method as for the isomerically defined glycosylated asparagine standards, the linkage variables were determined for the different glycopeptides carrying a N4H5S2 glycan. The linkage variable ranged from 0.20 to 0.25 across the glycopeptides, with the linkage variable of NKT being highest at 0.25 on average and lowest being NK at 17%. All other peptides showed linkage variables between 0.21 and 0.24 (Figure 6A). These linkage variables corresponded with almost exclusively (90–100%) α 2,6-linked sialic acids (Figure 5), which was in accordance with the literature.^{28,29} The similarity of the linkage variables indicated that the presence of amino acids A, K, V and T did not directly affect the linkage variable to a large degree.

Next to SGP, trastuzumab was chosen as a glycoprotein to assess the effect of the peptide backbone on the linkage variable. To this end, the protein was first digested by trypsin and, subsequently, split into two fractions. One of these fractions was defucosylated overnight, while both fractions were further partially digested with Pronase to obtain a range of glycopeptides of different sizes. As far as we could see, core fucosylation had no apparent effect on the linkage variable, whereas the peptide appeared to have a minor effect (Figure 6B,C). IgG N4H4(F1) is known to exist with both 3- and 6-branched galactosylation, although the ratio on trastuzumab in particular has not previously been reported. It is conceivable that an interaction exists between the galactose position and the digestion efficiency of particular amino acids, but this would require substantial further investigation.^{30,31}

DISCUSSION

While MS is, nowadays, one of the primary methods for glycan and glycoprotein analysis, its dependence on mass means that the important isomeric characteristics of glycosylation are difficult to access. By combining several oxonium ions across collision energies into a single “linkage variable”, we were able to discriminate isomer ratio differences as low as 10% on a variety of peptides. We could observe a consistent linkage variable across varying peptide lengths of sialylglycopeptide and trastuzumab, although it should be noted that the glycopeptides used in our experiments were relatively small and did not include all amino acids.

Interestingly, from our data, it appears that the saccharides connected to the 3-branch of the glycan are generally fragmented at lower collision energies than the saccharide

groups connected to the 6-branch of the glycan. A possible explanation for this consistent HCD fragmentation behavior is that linkages to the 3-carbon are more rigid than linkages to the 6-carbon, as the latter has more axes of rotation. This seemed to be the case as well for the α 2,3-linked NeuAc residues, which were more susceptible to dissociation than their α 2,6-linked counterparts. It would be interesting to assess whether this is only the case for the structures we have investigated or whether it proves to be a general rule that can be applied to future unknown glycopeptides. Currently, the main limiting factor for which isomeric properties can be analyzed is the availability of isomerically defined glycosylated asparagine standards. General fragmentation rules, such as the 3-branch always fragmenting before the 6-branch, could help in distinguishing future unknown samples. However, for relative quantification of glycopeptide, isomer mixtures still need to be defined and analyzed.

As it is, the method allows determination of the glycan structure of a glycopeptide without having peptide-specific parameters,¹⁴ apart from *m/z* ratio, intensity and retention time. In principle, this should enable easy incorporation into standard proteomics workflows, opening up a wide range of possibilities. For instance, it may be possible to generate a more detailed structural overview of antibody glycosylation in plasma samples and thereby a better understanding of their differing effector functions and half-life characteristics.³² The method could also be instrumental in comparing normal human protein glycosylation to its alternative made by cell systems or even tumor cells; making the structural elements of glycosylation accessible would open up a new dimension of potential biomarker research.³³

For now, however, the method has not yet been tested in complex samples, which are expected to require more complex data curation. On top of that, a complicating factor for system-wide glycoproteomics may be glycan-structure-specific proteolysis; it could very well be that certain proteoglycoforms are more easily cleaved by proteases than others, particularly when structural elements of glycosylation come into play. In addition, for purposes of biological investigation, it will be important to expand the repertoire of distinguishable glycan characteristics into directions such as α - and β -linkage, fucose position and linkage, and antennary GlcNAc position. One way to achieve this may be the combination of our workflow with contemporary separation methods such as ion mobility.

As it is, we have developed a glycoproteomics-compatible method for distinguishing between glycopeptide isomers for branching position and sialic acid linkage. The method is largely independent of the connected peptide and is expected to allow for the first instances of structural glycoproteomics by use of MS/MS alone.

ASSOCIATED CONTENT

Data Availability Statement

Data associated with the manuscript has been deposited within MassIVE repository “MSV000091172” and can additionally be accessed via <ftp://massive.ucsd.edu/MSV000091172/>.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.3c01319>.

Table S1: List of oxonium ions for MS/MS integration;

Table S2: Overview of calculations to construct linkage variables for isomer determination; Table S3. Overview

of p-values obtained when performing a student's *t* test between 10% isomer mixture steps; Supporting Information, 1: Python-script for data analysis; Supporting Information, 2: R-script for data analysis; Figures S1–S9: Additional details on the fragmentation differences between glycopeptides isomers (PDF)

AUTHOR INFORMATION

Corresponding Author

Karli R. Reiding – Biomolecular Mass Spectrometry and Proteomics, Utrecht Institute for Pharmaceutical Sciences and Bijvoet Center for Biomolecular Research, University of Utrecht, Utrecht 3584 CH, The Netherlands; Netherlands Proteomics Center, Utrecht 3584 CH, The Netherlands; Email: k.r.reiding@uu.nl

Authors

Joshua C.L. Maliepaard – Biomolecular Mass Spectrometry and Proteomics, Utrecht Institute for Pharmaceutical Sciences and Bijvoet Center for Biomolecular Research, University of Utrecht, Utrecht 3584 CH, The Netherlands; Netherlands Proteomics Center, Utrecht 3584 CH, The Netherlands; orcid.org/0009-0004-3748-1570

J. Mirjam A. Damen – Biomolecular Mass Spectrometry and Proteomics, Utrecht Institute for Pharmaceutical Sciences and Bijvoet Center for Biomolecular Research, University of Utrecht, Utrecht 3584 CH, The Netherlands; Netherlands Proteomics Center, Utrecht 3584 CH, The Netherlands

Geert-Jan P.H. Boons – Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences and Bijvoet Center for Biomolecular Research, University of Utrecht, Utrecht 3584 CG, The Netherlands; Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30602, United States; Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia 30602, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.analchem.3c01319>

Author Contributions

J.C.L.M. performed experiments and data analysis and wrote the manuscript. G.-J.P.H.B. provided isomerically defined glycosylated asparagine standards. J.M.A.D. performed experiments. K.R.R. performed experiments and data analysis, wrote the manuscript, designed the project, and contributed the funding. All authors reviewed and edited the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Roche Diagnostics for kindly providing Pronase and tastuzumab, Dario Cramer for helpful initial digestion experiments, Hillary Hoppenbrouwers for her assistance with setting up cotton-HILIC-based SGP extraction from egg yolk powder, and Gerlof Bosman for his efforts on glycopeptide synthesis. Furthermore, we would like to thank Javier Sastre Toraño for providing the isomerically defined glycosylated asparagine standards. This project was funded by the Dutch Research Council (NWO) Project VI.Veni.192.058, awarded to K.R.R. G.-J.P.H.B. acknowledges the ERC Advanced Grant 101020769.

REFERENCES

- (1) Varki, V.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; et al. *Essentials of Glycobiology* [Internet]; Cold Spring Harbor Laboratory Press, 2022; DOI: [10.1101/9781621824213](https://doi.org/10.1101/9781621824213).
- (2) An, H. J.; Froehlich, J. W.; Lebrilla, C. B. *Curr. Opin. Chem. Biol.* **2009**, *13* (4), 421–426.
- (3) Chou, H. H.; Takematsu, H.; Diaz, S.; Iber, J.; Nickerson, E.; Wright, K. L.; Muchmore, E. A.; Nelson, D. L.; Warren, S. T.; Varki, A. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (20), 11751–11756.
- (4) Stevens, J.; Blixt, O.; Paulson, J. C.; Wilson, I. A. *Nature Reviews Microbiology* **2006**, *4* (11), 857–864.
- (5) Ito, T.; Couceiro, J. N. S. S.; Kelm, S.; Baum, L. G.; Krauss, S.; Castrucci, M. R.; Donatelli, I.; Kida, H.; Paulson, J. C.; Webster, R. G.; Kawaoka, Y. *J. Virol.* **1998**, *72* (9), 7367–7373.
- (6) Reiding, K. R.; Bondt, A.; Franc, V.; Heck, A. J. R. *TrAC Trends in Analytical Chemistry* **2018**, *108*, 260–268.
- (7) Chen, R.; Stupak, J.; Williamson, S.; Twine, S. M.; Li, J. *Rapid Commun. Mass Spectrom.* **2019**, *33* (15), 1240–1247.
- (8) Zhu, R.; Huang, Y.; Zhao, J.; Zhong, J.; Mechref, Y. *Anal. Chem.* **2020**, *92* (14), 9556–9565.
- (9) Van Der Burgt, Y. E. M.; Siliakus, K. M.; Cobbaert, C. M.; Ruhaak, L. R. *J. Proteome Res.* **2020**, *19* (7), 2708–2716.
- (10) Huang, Y.; Nie, Y.; Boyes, B.; Orlando, R. J. *Biomol. Tech.* **2016**, *27* (3), 98–104.
- (11) Kammeijer, G. S. M.; Jansen, B. C.; Kohler, I.; Heemskerk, A. A. M.; Mayboroda, O. A.; Hensbergen, P. J.; Schappler, J.; Wührer, M. *Sci. Rep.* **2017**, *7* (1), 1–10.
- (12) de Haan, N.; Yang, S.; Cipollo, J.; Wührer, M. *Nature Reviews Chemistry* **2020**, *4* (5), 229–242.
- (13) Ács, A.; Ozohanics, O.; Vékey, K.; Drahoš, L.; Turiák, L. *Anal. Chem.* **2018**, *90* (21), 12776–12782.
- (14) Pett, C.; Nasir, W.; Sihlbom, C.; Olsson, B. M.; Caixeta, V.; Schorlemer, M.; Zahedi, R. P.; Larson, G.; Nilsson, J.; Westerlind, U. *Angew. Chem., Int. Ed. Engl.* **2018**, *57* (30), 9320–9324.
- (15) Wu, R.; Chen, X.; Wu, W. J.; Wang, Z.; Hung, Y. L. W.; Wong, H. T.; Chan, T. W. D. *Rapid Commun. Mass Spectrom.* **2020**, *34* (9), No. e8751.
- (16) Guttman, M.; Lee, K. K. *Anal. Chem.* **2016**, *88* (10), 5212–5217.
- (17) Both, P.; Green, A. P.; Gray, C. J.; Šardžik, R.; Voglmeir, J.; Fontana, C.; Austeri, M.; Rejzek, M.; Richardson, D.; Field, R. A.; Widmalm, G.; Flitsch, S. L.; Evers, C. E. *Nat. Chem.* **2014**, *6* (1), 65–74.
- (18) Wong, T. L.; Mooney, B. P.; Cavallero, G. J.; Guan, M.; Li, L.; Zaia, J.; Wan, X. F. *J. Proteome Res.* **2023**, *22* (1), 62–77.
- (19) Feng, X.; Shu, H.; Zhang, S.; Peng, Y.; Zhang, L.; Cao, X.; Wei, L.; Lu, H. *Anal. Chem.* **2021**, *93* (47), 15617–15625.
- (20) Barroso, A.; Giménez, E.; Konijnenberg, A.; Sancho, J.; Sanz-Nebot, V.; Sobott, F. *J. Proteomics* **2018**, *173*, 22–31.
- (21) Liu, L.; Prudden, A. R.; Capicciotti, C. J.; Bosman, G. P.; Yang, J. Y.; Chapla, D. G.; Moremen, K. W.; Boons, G. J. *Nat. Chem.* **2019**, *11* (2), 161–169.
- (22) Broszeit, F.; van Beek, R. J.; Unione, L.; Bestebroer, T. M.; Chapla, D.; Yang, J. Y.; Moremen, K. W.; Herfst, S.; Fouchier, R. A. M.; de Vries, R. P.; Boons, G. J. *Nat. Commun.* **2021**, *12* (1), 1–12.
- (23) Broszeit, F.; Tzarum, N.; Zhu, X.; Nemanichvili, N.; Eggink, D.; Leenders, T.; Li, Z.; Liu, L.; Wolfert, M. A.; Papanikolaou, A.; Martínez-Romero, C.; Gagarinov, I. A.; Yu, W.; García-Sastre, A.; Wennekes, T.; Okamatsu, M.; Verheije, M. H.; Wilson, I. A.; Boons, G. J.; de Vries, R. P. *Cell Rep* **2019**, *27* (11), 3284–3294.
- (24) Gagarinov, I. A.; Li, T.; Toraño, J. S.; Caval, T.; Srivastava, A. D.; Kruijtz, J. A. W.; Heck, A. J. R.; Boons, G. J. *J. Am. Chem. Soc.* **2017**, *139* (2), 1011–1018.
- (25) Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M.; Wührer, M. *Anal. Chem.* **2011**, *83* (7), 2492–2499.
- (26) Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M. *J. Proteome Res.* **2008**, *7* (4), 1650–1659.
- (27) Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.;

Prestegard, J. J.; Schnaar, R. L.; Freeze, H. H.; Marth, J. D.; Bertozzi, C. R.; Etzler, M. E.; Frank, M.; Vliegthart, J. F. G.; Lütke, T.; Perez, S.; Bolton, E.; Rudd, P.; Paulson, J.; Kanehisa, M.; Toukach, P.; Aoki-Kinoshita, K. F.; Dell, A.; Narimatsu, H.; York, W.; Taniguchi, N.; Kornfeld, S. *Glycobiology* **2015**, *25* (12), 1323–1324.

(28) Seko, A.; Koketsu, M.; Nishizono, M.; Enoki, Y.; Ibrahim, H. R.; Juneja, L. R.; Kim, M.; Yamamoto, T. *Biochim. Biophys. Acta* **1997**, *1335* (1–2), 23–32.

(29) Liu, L.; Prudden, A. R.; Bosman, G. P.; Boons, G. J. *Carbohydr. Res.* **2017**, *452*, 122–128.

(30) Sanchez-De Melo, I.; Grassi, P.; Ochoa, F.; Bolivar, J.; García-Cózar, F. J.; Durán-Ruiz, M. C. *J. Proteomics* **2015**, *127*, 225–233.

(31) Falck, D.; Jansen, B. C.; Plomp, R.; Reusch, D.; Habberger, M.; Wührer, M. *J. Proteome Res.* **2015**, *14* (9), 4019–4028.

(32) Anthony, R. M.; Nimmerjahn, F.; Ashline, D. J.; Reinhold, V. N.; Paulson, J. C.; Ravetch, J. V. *Science* **2008**, *320* (5874), 373–376.

(33) Croset, A.; Delafosse, L.; Gaudry, J. P.; Arod, C.; Glez, L.; Losberger, C.; Begue, D.; Krstanovic, A.; Robert, F.; Vilbois, F.; Chevalet, L.; Antonsson, B. *J. Biotechnol.* **2012**, *161* (3), 336–348.